On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge

Barbara Citko

This article argues in favor of a new type of Merge, Parallel Merge, which combines the properties of External Merge and Internal Merge. Parallel Merge creates symmetric, multidominant structures, which become antisymmetric in the course of the derivation. The main empirical goal of the article is to revive a multidominance approach to across-the-board *wh*-questions and to show that a number of otherwise puzzling properties of across-the-board questions follow naturally from such an account.

1 Parallel Merge

The main theoretical goal of this article is to argue for a new type of Merge, called *Parallel Merge*, whose existence is a natural consequence of Chomsky's (2001) view of phrase structure and movement. Parallel Merge creates symmetric, multidominant structures, which become antisymmetric in the course of the derivation. The main empirical goal is to provide new evidence in favor of a Parallel Merge approach to across-the-board (ATB) *wh*-questions and to show that a number of otherwise puzzling properties of ATB questions follow naturally from such an account.

Chomsky (2001) distinguishes between two types of Merge: *Internal Merge* and *External Merge*. External Merge, the ''canonical'' type of Merge, takes two distinct rooted structures and joins them into one.

(1) *External Merge*

$$\alpha \qquad \beta \qquad \rightarrow \qquad \overset{\displaystyle \wedge}{\underset{\alpha \qquad \beta}{}}$$
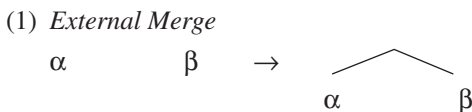
Internal Merge differs from External Merge in that it takes a subpart of an existing structure as one of the two objects. Internal Merge thus yields the effects of syntactic movement. The displaced element, such as β in (2), instead of being copied and merged into a new position, is simply
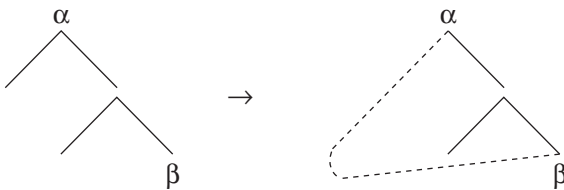
remerged/internally merged into its new position.[1] In what follows, I will represent Internal Merge with a dashed line.

(2) a. *Copy theory of movement* (Chomsky 1995)

α   α

                →   β

                β   β
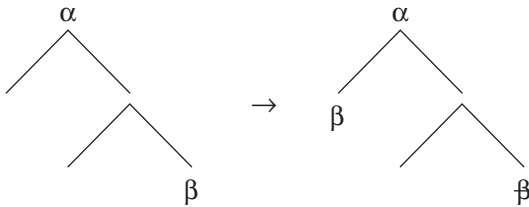
b. *Internal Merge theory of movement* (Epstein et al. 1998, Gärtner 1999, Chomsky 2001, Starke 2001, Zhang 2002)

α   α

                →

                β   β

The existence of External Merge and Internal Merge predicts the existence of a third type, combining the properties of both. This third type, which I will refer to as *Parallel Merge*, is like External Merge in that it involves two distinct rooted objects (α and β), but it is like Internal Merge in that it combines the two by taking a subpart of one of them, as shown in (3).[2]

(3) *Parallel Merge*

α   β   →   α   β

α   γ   α   γ   β

---
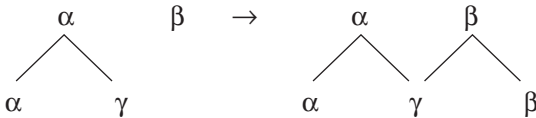
[1] The Internal Merge theory of movement must be augmented with a proposal regarding the pronunciation and interpretation of the remerged element. For example, in cases of overt movement, it will be pronounced in the remerged position. An interesting issue that I will leave open for the time being is whether the pronunciation and interpretation sites can be derived from independently motivated principles, such as economy or feature checking. This is essentially what Nunes (1995), who works within the copy theory of movement, does to determine which copy is deleted and which one is pronounced.

[2] The Parallel Merge operation proposed here bears some resemblance to Chomsky's (2001) Pair Merge operation, which is involved in the derivation of adjuncts. Both create multidominant (or multiplanar) objects. The two proposals differ, however, with respect to how such multidominant objects are linearized. Chomsky derives the linear ordering of adjuncts from an operation Simplify, which converts Pair Merge objects to standard Set Merge objects. As we will see in the next section, Parallel Merge does not require any special mechanism to linearize multidominant objects.

Another logical possibility, brought to my attention by an anonymous reviewer, is that Parallel Merge targets subparts of two distinct objects. I do not see any conceptual reasons to exclude this possibility. For the purposes of this article, however, I will limit my attention to the kind of Parallel Merge illustrated in (3).

Since the existence of Parallel Merge structures is a theoretical possibility, it is important to ask whether there are any empirical arguments against it. A potential issue, which I will argue is only apparent, involves linearization. More specifically, the issue is how Parallel Merge structures are mapped onto a linear string. This is the topic of the next section, where I develop a way to make Parallel Merge structures compatible with perhaps the most restrictive proposal regarding the mapping between syntax and phonology, namely, Kayne's (1994) antisymmetry proposal.

2 Parallel Merge and Linearization

Parallel Merge structures are clearly incompatible with Kayne's (1994) Linear Correspondence Axiom (LCA), which derives linear precedence from strict asymmetric c-command, thus banning symmetric structures from syntax. An interesting issue that arises here, also discussed by Chomsky (1995), concerns the status of the LCA in a bare phrase structure theory, which I adopt here. Kayne's original formulation of the LCA relies on the existence of intermediate nonbranching levels, which are banned in a bare phrase structure theory. The structures that the LCA applies to, such as (4a), simply cannot be generated. What results instead is (4b).

(4) a.
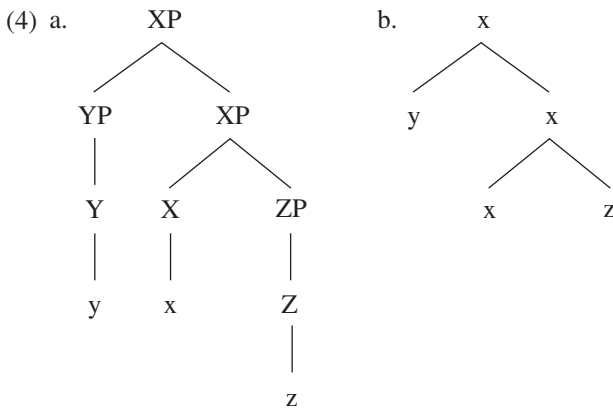
However, the basic insight of the LCA—namely, the intuition that linear ordering is part of the phonological component and can be unambiguously derived from asymmetric c-command—is fully compatible with bare phrase structure. Chomsky (1995:334–340) discusses precisely this issue and offers a ''bare'' reformulation of the LCA.[3]

---

[3] A related question, brought to my attention by an anonymous reviewer, is how the LCA works in a theory that replaces Move with Internal Merge. Again, it seems reasonable to believe that the basic insights of the LCA can be stated in such a system. All that is needed is to assume that for elements that have been both externally and internally merged, it is either the Internal Merge position or the External Merge position that will count for the purposes of the LCA.

BARBARA CITKO

The details of such a reformulation are not directly relevant to the present discussion.[4] What is directly relevant is that the LCA fails to unambiguously and totally order Parallel Merge structures of the kind given in (5).

(5) x^{max} y^{max}

     x z^{max} y
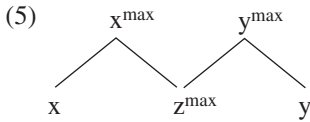
Irrespective of how x^{max} and y^{max} will ultimately end up ordered with respect to each other, there is a problem with z^{max}. For the sake of concreteness, let us assume that at some later stage in the derivation, y^{max} will end up c-commanding x^{max}. This means that the terminals of y^{max} will end up preceding the terminals of x^{max}. Since z^{max} is part of both x^{max} and y^{max}, it will be subject to contradictory linearization requirements; it will end up both preceding and following itself.

The solution to this problem lies in determining where exactly the LCA applies. Chomsky (1995), contra Kayne (1994), suggests that there is no reason for the LCA to order an element that will disappear at Spell-Out, such as a trace or an unpronounced copy.[5] In other words, since the LCA is linked to pronunciation, there is no reason why it should have to apply *throughout* the entire derivation. Chomsky's version of the LCA thus allows Parallel Merge structures as long as the shared element undergoes overt movement (more accurately, Internal Merge) into a higher position. In effect, this theory allows Parallel Merge as long as its effects are invisible at Spell-Out.[6] This approach to the LCA is also the basic insight of Moro's (2000) *Dynamic Antisymmetry* proposal, which allows symmetric structures, ''points of symmetry'' in Moro's terminology, before Spell-Out. Even though Moro, who focuses on small clauses, multiple specifiers, and clitic adjunction structures, does not include multidominant structures in the range of permissible symmetric structures, his theory certainly allows them.[7]

---

[4] Both the original formulation of the LCA and Chomsky's ''bare'' reformulation derive specifier-head-complement ordering only for branching complements. In (4b), neither x nor z asymmetrically c-commands the other. Kayne (1994) bans such structures altogether; Chomsky (1995) allows them only if one of the offending elements (either x or y) moves out of the shared structure.

[5] The question, brought to my attention by one of the reviewers, is whether the assumption that traces are subject to the LCA is crucial for Kayne (1994). Kayne notes the possibility that traces might be invisible to the LCA. However, he then concludes that the issue of their invisibility does not arise on the copy theory of movement (Kayne 1994:chap. 2, n. 3).

[6] This opens up the possibility that Parallel Merge is possible also after Spell-Out, where linearization is no longer an issue. See Epstein et al. 1998 for a suggestion along these lines.

[7] My proposal differs from Moro's in how it handles movement. For Moro (2000), who wants to eliminate feature-driven movement from the grammar, the need for antisymmetry is the sole motivation for movement. I maintain a more ''conservative'' approach, on which movement is driven by uninterpretable formal features, and antisymmetry is its consequence rather than its driving force. Even in Moro's Dynamic Antisymmetry system, there is still a need for some form of feature-driven movement. The search for antisymmetry will determine why things have to move; it will not determine where they are going to land. Thus, features on the target of movement (such as the EPP-feature on T in the case of subject raising) will need to be involved in determining the landing site for moved elements.

The rest of the article focuses on the empirical side of Parallel Merge. It develops a Parallel Merge account of ATB *wh*-questions, questions involving what looks like simultaneous extraction of a *wh*-element from two (or more) conjuncts.

(6) What did John recommend and Mary read?

It shows that three otherwise rather striking properties of such questions follow naturally from a Parallel Merge approach: (a) the presence of the so-called matching effects, the requirement that a *wh*-pronoun simultaneously satisfy case requirements imposed by two distinct verbs (Dyła 1984, Franks 1993, 1995), (b) the lack of covert ATB *wh*-movement (Cho and Zhou 1999, Wu 1999, Bošković and Franks 2000, Citko and Grohmann 2000), and (c) the lack of ATB *wh*-questions with multiple fronted *wh*-words.

3 A Parallel Merge Approach to Across-the-Board *Wh*-Questions

The idea that ATB *wh*-questions involve parallel structures goes back to Williams 1978, where they are analyzed as involving parallel phrase markers, ''factors'' in Williams's terminology (see also Goodall 1983, 1987, Muadz 1991, Moltmann 1992).

On the Parallel Merge view, the parallelism is captured in a different way, by allowing a single *wh*-phrase to merge with elements inside both conjuncts. The derivation of the embedded ATB *wh*-question in (7), for example, involves merging the *wh*-pronoun *what* with the verb *read* and subsequently parallel-merging the verb *recommended* with it, as shown in (8a–b).

(7) I wonder *what Gretel recommended and Hansel read.*

(8) a. *Merge* read *and* what, *project* read

$$V^{max}$$

read   what
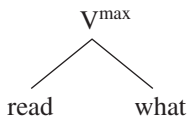
b. *Parallel-merge* recommended *and* what, *project* recommended

$$V^{max} \qquad V^{max}$$

read   what   recommended
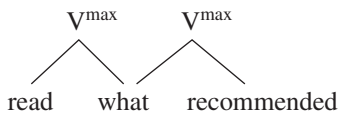
The ''flat'' tree in (8b) does not portray linear order. A more perspicuous diagram is given in (9).

(9) V^{max} V^{max}

read   recommended   what
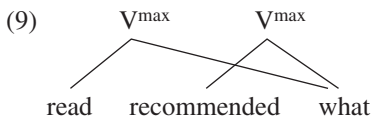
From this point on, the derivation proceeds in parallel until the two clauses given in (10) have been constructed. For the sake of clarity, the following discussion ignores base-generated positions of the subjects.

(10)

```
        Tᵐᵃˣ                              Tᵐᵃˣ
       /    \                            /    \
   Hansel    T                       Gretel    T
            /  \                             /   \
           T    vᵐᵃˣ                        T     vᵐᵃˣ
               /   \                             /    \
              v     Vᵐᵃˣ                        v      Vᵐᵃˣ
                   /                                  /  \
                 read              recommended    what
```
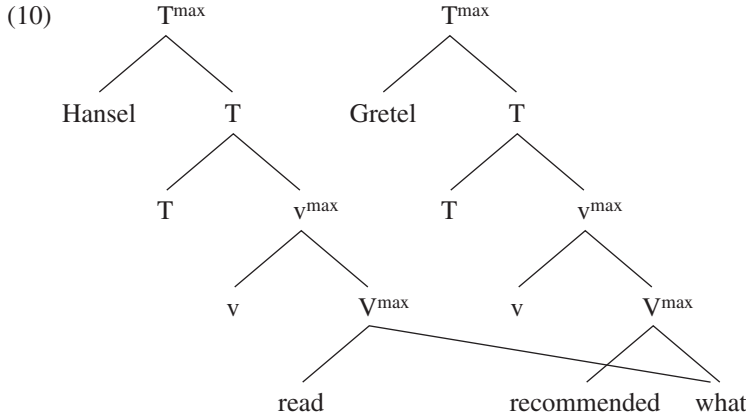
An interesting question that Parallel Merge raises involves Case checking: namely, how a single element can check the Case features of two distinct heads. The question does not arise if we assume the probe-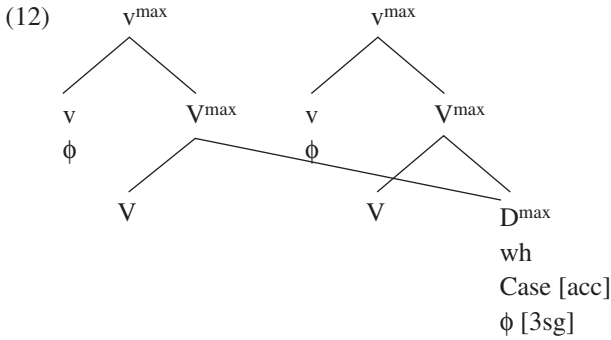goal Case system developed in Chomsky 2000, 2001, which replaces feature checking with feature valuation and treats Case not as a primitive feature but as a reflex of φ-features on T and v heads. In this system, uninterpretable features enter the derivation unvalued. In (11), for example, both v heads contain unvalued φ-features and the nominal *what* contains an unvalued Case feature.[8]

(11)

```
        vᵐᵃˣ                              vᵐᵃˣ
       /    \                            /    \
      v      Vᵐᵃˣ                       v      Vᵐᵃˣ
    φ [u]   /    \                    φ [u]   /  \
           V      V        Dᵐᵃˣ
                          wh
                          Case [u]
                          φ [3sg]
```
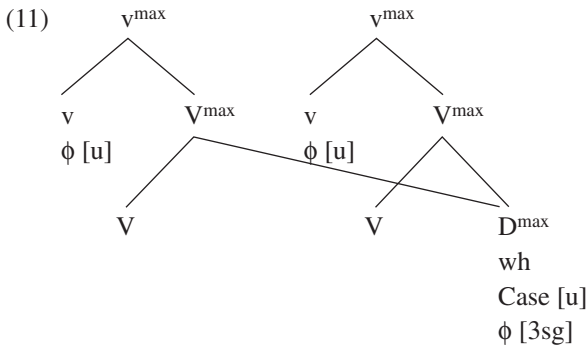
The Agree operation provides values to unvalued features under appropriate structural conditions, as shown in (12).

---

[8] Borrowing the convention from Pesetsky and Torrego (2001), I use the symbol [u] to represent unvalued features.

(12)

```
        vᵐᵃˣ                           vᵐᵃˣ
       /    \                         /    \
      v     Vᵐᵃˣ                     v     Vᵐᵃˣ
      φ    /    \                    φ    /    \
          V      \                       V      \
                  \                              \
                   V        ────────────        Dᵐᵃˣ
                                                 wh
                                                 Case [acc]
                                                 φ [3sg]
```

The probe-goal system does not exclude the possibility that a single goal can agree simultane-ously with two (or more) probes. Since Case valuation is divorced from movement, nothing pre-vents a single element from entering into multiple Case valuation relationships.[9]

This view of Case checking sets the present proposal apart from a conceptually quite similar account of ATB movement, namely, Nunes's (1995) sideward movement account. Both accounts involve a single *wh*-pronoun in an ATB question. Under Nunes's account, this *wh*-pronoun first merges into its θ-position inside the first conjunct and then moves to the other conjunct in a sideward (rather than upward) fashion, before moving to Spec,C. Nunes has to assume that the Case features of the *wh*-pronoun remain active after it has undergone Agree with the v head inside the first conjunct. Otherwise, there would be no way for the v head inside the other conjunct to have its features valued. It is not clear to me how the Case features of the *wh*-pronoun ''know'' when to remain active after one Agree operation has taken place. They have to remain active if there is another Agree operation coming up. On the Parallel Merge view proposed here, the two Agree operations are simultaneous, which allows the shared DP to be active for both probes simultaneously without any look-ahead property.

The next two steps in the derivation involve merging the two clauses with the conjunction head. As a result, one becomes its complement, and the other its specifier, as shown in (13).

---

[9] This can be thought of as the conceptual reverse of Hiraiwa's (2001) Multiple Agree mechanism, which allows a single probe to agree simultaneously with two goals. This is what happens, for example, when a single interrogative complementizer agrees simultaneously with features of two *wh*-phrases in its c-command domain.

(13)
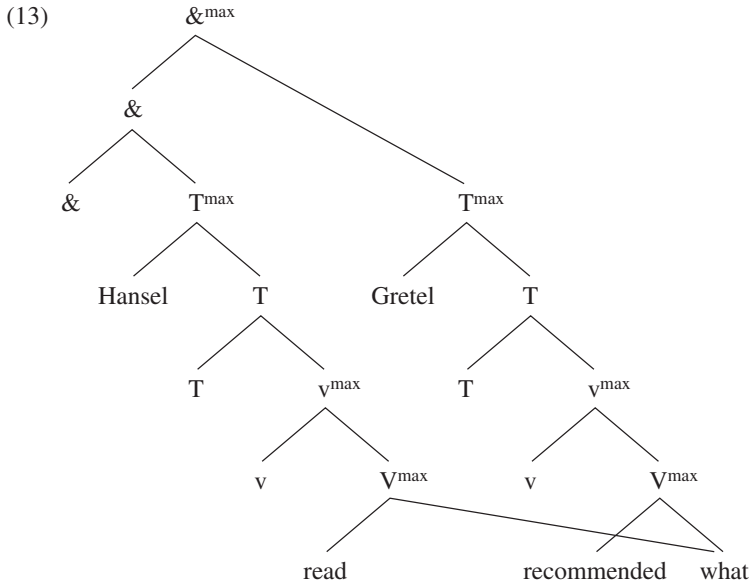
This structure assumes a fairly standard approach to coordination, in which the two conjuncts are in an asymmetric relationship with respect to each other and the conjunction *and* heads the entire projection. Evidence favoring such an approach comes from variable binding (Munn 1993), extraposition (Collins 1988, Munn 1993), and pro-form replacement (Zoerner 1995).[10] Thus, the

[10] For the purposes of this article, I will not try to choose among the various asymmetric approaches to coordination proposed in Munn 1993 (i), Progovac 1998 (ii), and Collins 1988 (iii), all of which are compatible with my proposal.
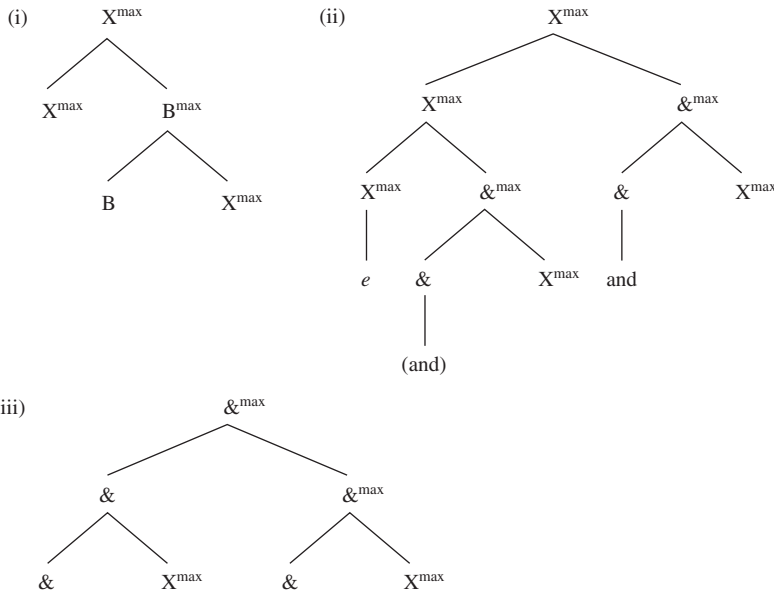
quantifier in the first conjunct can bind a variable inside the second conjunct but not vice versa, as shown in (14). Furthermore, the conjunction head forms a constituent with the second conjunct, as shown by the extraposition data in (15) and by the *etc.* pro-form replacement data in (16).

(14) a. *Every professor* and *his spouse* attended the convocation.
    b. *His spouse and *every professor* attended the convocation.

(15) a. John read a book yesterday, and the newspapers.
    b. *John read the newspapers yesterday, the book and.

(16) a. I bought jam, bread, etc.
    b. *I bought jam, bread, and etc.

Next, the conjunction phrase merges with a complementizer head, and the complementizer head undergoes Internal Merge with the *wh*-pronoun. I assume that internally merged elements can be spelled out only once. In the case at hand, the *wh*-pronoun will be spelled out in Spec,C.

(17) *Remerge (internal-merge)* what *with C*
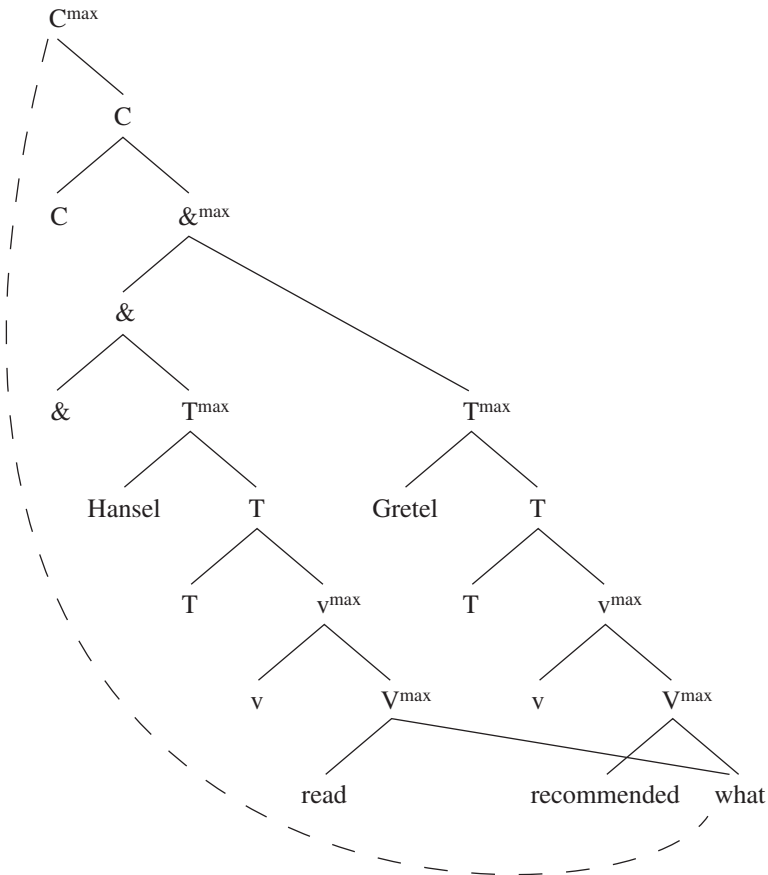
Note that the structure in (17) still contains a shared element, *what*. However, since the shared element has undergone Internal Merge into a higher position, which is the position where it will be pronounced, as far as the LCA is concerned, the structure in (17) is equivalent to the more ''standard'' one in (18).[11]
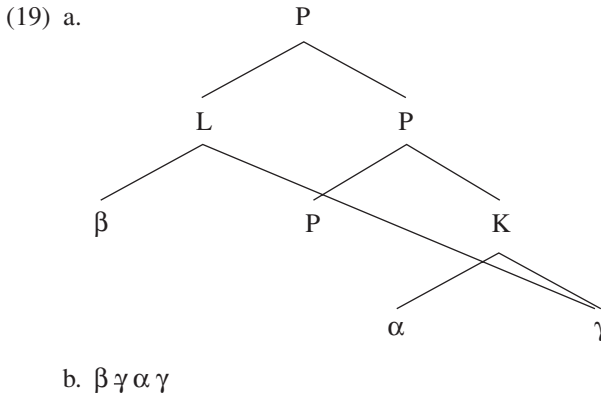
(18)

C^{max}

$what_i$ C

   C   &

     T^{max} &

   Gretel recom-   and T^{max}
   mended e_i

   Hansel
   read e_i
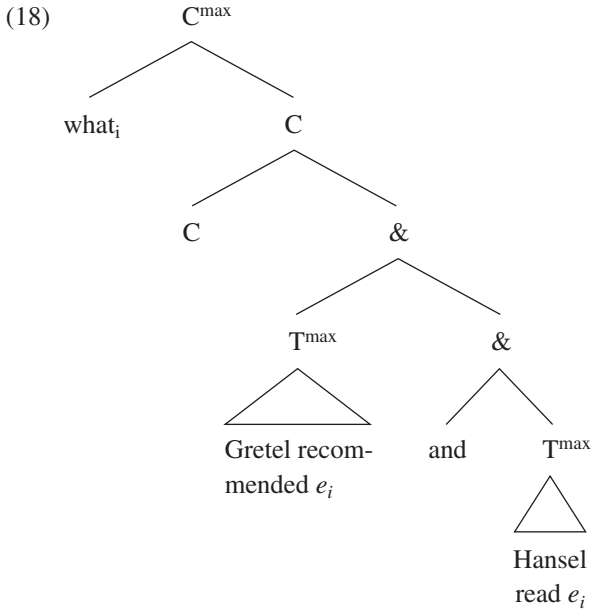
    The view that linearization is a consequence of movement sets the present proposal apart from many other multidominance proposals, which either do not address linearization at all, or enrich the grammar with rules designed specifically to handle linearization of multidominant structures. It is close in spirit to Wilder's (1999) analysis of right node raising, which also explores the interaction of multidominant structures with the LCA. The specifics of the two accounts are quite different, however. On the Parallel Merge view advocated here, movement is what makes linearization possible. On Wilder's (1999) view, multidominant structures can be linearized without movement of the shared constituent. Empirically, the two proposals make different predictions regarding linearization of multidominant structures of the type shown in (19a).

---

[11] A reviewer inquires whether other processes besides movement could be involved in linearizing Parallel Merge structures. Ellipsis is a plausible candidate for such a process. I will leave further investigation of this intriguing possibility for future research.

(19) a.

                        P
                    ╱   ╲
                  L   P
               ╱   ╲   ╱   ╲
             β   P   K
                   ╱   ╲   ╱
                 α   γ

    b. β γ α γ

On Wilder's account, (19a) will be linearized as (19b). The shared element, γ, will be pronounced
in the final conjunct, and there will be a gap (a copy of γ) in the nonfinal conjunct. On the Parallel
Merge account, (19a) is not linearizable at all because the shared element γ has not moved out
of the shared structure (or been remerged in a nonshared position). As we will see in the next
section, the properties of ATB *wh*-questions in *wh*-in-situ languages follow precisely because we
posit that such structures are not linearizable.

4 Consequences of the Parallel Merge Approach to Across-the-Board *Wh*-Questions

4.1 Matching Effects

One of the most straightforward consequences of the Parallel Merge approach to ATB questions
comes from the so-called matching effects, the requirement that the fronted *wh*-pronoun match
in case the gaps inside the two conjuncts (Borsley 1983, Dyła 1984, Franks 1995).

(20) [$_{CP}$ WH$_{\alpha Case}$ [$_{TP}$ . . . $e_{\alpha Case}$ . . .] and [$_{TP}$ . . . $e_{\alpha Case}$ . . .]]

The contrast between the Polish examples in (21a) and (21b) illustrates the case-matching require-
ment.

(21) a. Kogo   Jan lubi *e*   a   Maria podziwia *e*?
            who.ACC Jan likes *e*.ACC and Maria admires   *e*.ACC
            'Who does Jan like and Maria admire?'
      b. *Kogo/Komu Jan lubi *e*   a   Maria ufa   *e*?
            who.ACC/DAT Jan likes *e*.ACC and Maria trusts *e*.DAT
            'Who does Jan like and Maria trust?'

In the grammatical example (21a), the fronted accusative *wh*-pronoun *kogo* 'who' matches the
gaps inside the two conjuncts, satisfying the case requirements of both verbs. Example (21b), on
the other hand, is ungrammatical because the verbs inside the two conjuncts differ in their case
requirements: the verb *lubić* 'like' requires an accusative object, whereas the verb *ufać* 'trust'
requires a dative object. Therefore, neither the accusative *wh*-pronoun *kogo* nor the dative *wh*-

pronoun *komu* can match both gaps simultaneously. The Parallel Merge mechanism allows for a very simple account of this contrast in grammaticality. Since the *wh*-pronoun is merged with two verbs, it must simultaneously satisfy whatever category and case restrictions are imposed by both verbs. The derivation of the ungrammatical example (21b) would have to involve either of the structures given in (22). If the dative *wh*-pronoun *komu* undergoes Parallel Merge with the two verbs, the case requirements of the verb *lubi* 'likes' will remain unsatisfied (22a), and if the accusative *wh*-pronoun *kogo* undergoes Parallel Merge with the two verbs, the case requirements of the verb *ufa* 'trusts' will remain unsatisfied (22b).

(22) a. * V^{max} V^{max}

         lubi   ufa   komu$_{DAT}$

      b. * V^{max} V^{max}

         lubi   ufa   kogo$_{ACC}$
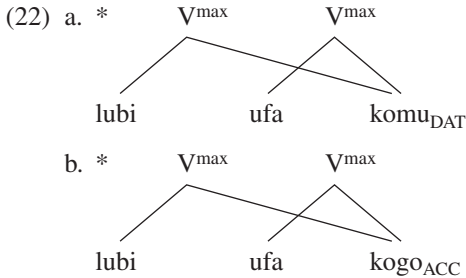
Interestingly, there is one exception to the case-matching requirement; mismatches in case are possible as long as there exists an appropriate syncretic form. The syncretic *wh*-forms of Polish are given in (23).

(23) *Case syncretism in Polish*

Nom	kto 'who'	*co* 'what'
Gen	*kogo*	czego
Dat	komu	czemu
Acc	*kogo*	*co*
Loc	*kim*	*czym*
Instr	*kim*	*czym*

The availability of an appropriate syncretic form is what accounts for the contrast in grammaticality between the examples in (24). Since there is no syncretic genitive/accusative *wh*-form in the inanimate series, the matching requirement cannot be satisfied in (24a). By contrast, since there is a syncretic genitive/accusative *wh*-form in the animate series, the matching requirement can be satisfied in (24b).[12]

---

[12] The ameliorating effects of syncretic forms are not limited to ATB questions. Coordinated Italian clitics show similar effects (Lori Repetti, pers. comm.).

(i) Mi   ha visto.
    me.ACC has seen
    'She/He saw me.'

(24) a. *Czego/Co   Jan nienawidzi *e*   a   Maria lubi *e*?
      whom.GEN/ACC Jan hates   *e*.GEN and Maria likes *e*.ACC
      'Who does Jan hate and Maria like?'

   b. Kogo   Jan nienawidzi *e*   a   Maria lubi *e*?
      who.ACC/GEN Jan hates   *e*.GEN and Maria likes *e*.ACC
   'Whom does Jan hate and Maria like?'

The ameliorating effects of case syncretism are a puzzle if we assume that the lexicon contains two *wh*-pronouns, one accusative and the other genitive, which happen to sound the same. On this view, it would simply be a mystery why inserting a genitive *wh*-pronoun as a complement of a verb that requires an accusative object, which normally results in ungrammaticality, should become grammatical if the two case forms happen to have the same phonological form. A more promising alternative is thus to assume that the lexicon contains a single *wh*-form, underspecified in such a way that it is compatible with both genitive and accusative case features.[13] For the sake of concreteness, I assume that lexical items are inserted postsyntactically during Spell-Out, following the Distributed Morphology framework (see Halle and Marantz 1993, Marantz 1995, among many others). In this framework, terminal syntactic nodes are purely abstract ''feature bundles'' with no phonological content. The example in (24b) thus involves the structure in (25) at Spell-Out.

 (ii) Mi   ha dato un bacio.
      me.DAT has given a   kiss
      'She/He gave me a kiss.'

(iii) Mi   ha visto e dato un bacio.
      me.ACC/DAT has seen and given a   kiss
      'She/He saw me and gave me a kiss.'

 (iv) *L'ha   visto e dato un bacio.
      him.ACC has seen and given a   kiss
      'She/He saw him and gave him a kiss.'

  (v) *Gli   ha visto e dato un bacio.
      him.DAT has seen and gave a   kiss
      'She/He saw him and gave him a kiss.'

 (vi) L'ha   visto.
      him.ACC has seen
      'She/He saw him.'

(vii) Gli   ha dato un bacio.
      him.DAT has given a   kiss
      'She/He gave him a kiss.'

[13] For other underspecification approaches to syncretism, see Jakobson 1958, Chvany 1986, Neidle 1988, Franks 1995.

(25)

$$v^{max} \qquad v^{max}$$

v
ϕ

V^{max}

v
ϕ

V^{max}

V

V

D^{max}
wh
Case [acc, gen]
ϕ [3sg]
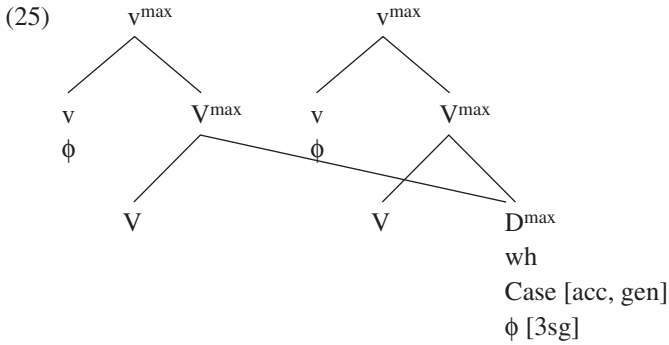
Since the lexicon contains a single form that is compatible with both accusative and genitive case features by virtue of underspecification, vocabulary insertion can proceed without any problems.[14] The ungrammatical example in (24a), on the other hand, involves the structure in (26) prior to vocabulary insertion.

(26)

$$v^{max} \qquad v^{max}$$

v
ϕ

V^{max}

v
ϕ

V^{max}

V

V

D^{max}
wh
animate
Case [acc, gen]
ϕ [3sg]
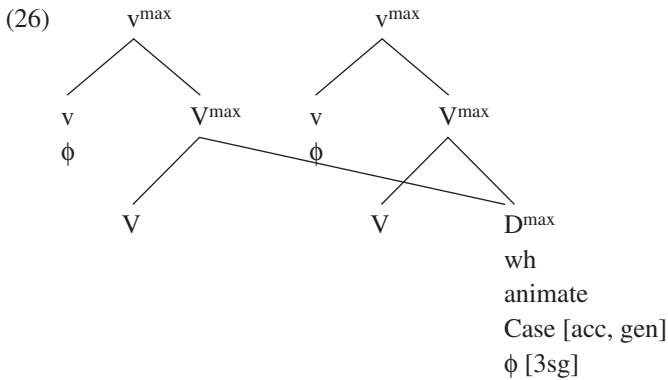
Since there is no single lexical item that can be inserted into this slot without a feature clash (there is no syncretic accusative/genitive form in the animate series), the result is ungrammatical.

4.2 Covert Across-the-Board Movement

Another interesting consequence of the Parallel Merge approach involves covert ATB movement. If movement is directly responsible for linearizing the structure, it must take place overtly. Covert ATB movement is thus predicted not to exist, since it takes place too late in the derivation to affect linearization.[15] This prediction is indeed borne out. Corroborating evidence comes from

---

[14] Late vocabulary insertion is not crucial to the present proposal. The ameliorating effects of case syncretism will also follow in a system with early vocabulary insertion, as long as the lexicon contains an appropriately underspecified element.

[15] An interesting prediction, brought to my attention by an anonymous reviewer, is that there could be cases of covert Parallel Merge, which could feed covert ATB movement. It seems, however, that recoverability considerations would substantially restrict covert Parallel Merge.

wh-in-situ languages, which nevertheless have overt ATB movement, and from the absence of ATB quantifier raising. Let us first consider ATB *wh*-questions in *wh*-in-situ languages (examples are from Cho and Zhou 1999; see also Wu 1999, Citko and Grohmann 2000).

(27) Zhangsan xihuan *shenme ren*   Lisi taoyan *shenme ren*?   (Chinese)
     Zhangsan like   which person Lisi hate   which person
     'Which person does Zhangsan like and which person does Lisi
     hate?'

(28) John-i   *enu   salam-ul*   cohaka-ko Mary-ka   *enu   salam-ul*   (Korean)
     John-NOM which person-ACC like-and   Mary-NOM which person-ACC
     miweha-ni?
     hate-Q
     'Which person does John like and which person does Mary hate?'

(29) John-ga   *dono hito-o*   aisitei-te Mary-ga   *dono hito-o*   (Japanese)
     John-NOM which person-ACC love-and Mary-NOM which person-ACC
     nikundeiru-no?
     hate-Q
     'Which person does John love and which person does Mary hate?'

These examples are not ungrammatical. However, they lack the reading typically associated with ATB questions, the so-called single-individual reading given in (30a), on which the question is about a single individual. They only allow a family-of-questions reading, given in (30b), on which the question is about two distinct individuals.[16]

(30) a. *which person x, Zhangsan/John likes x and Lisi/Mary hates x
     b.   which person x, Zhangsan/John likes x and which person y, Lisi/Mary hates y

For the single-individual reading to be available, the *wh*-phrase has to move overtly. This is rather puzzling since we are dealing with languages that are widely known to lack overt *wh*-movement.

(31) a. *Shenme ren*   Zhangsan xihuan Lisi taoyan?   (Chinese)
        which   person Zhangsan like   Lisi hate
        'Which person does Zhangsan like and Lisi hate?'

---

[16] Munn (1999) notes potential counterexamples to the generalization that ATB questions require single-individual answers.

 (i) Which man did Bill kill on Tuesday and Fred kill on Wednesday?
 (ii) Bill killed his first victim and Fred killed his second.

Such readings, however, involve functional *wh*-traces. Assuming the existence of functional traces yields pair-list readings without complicating (or abandoning) the ATB formalism. The pair-list reading depends on the sloppy identity of the argument index of the function.

 (iii) Which man^y did Bill$_x$ kill t_x^y on Tuesday and Fred$_x$ kill t_x^y on Wednesday?

   b. *Enu   salam-ul*   John-i   cohaka-ko Mary-ka   miweha-ni?   (Korean)
      which person-ACC John-NOM like-and   Mary-NOM hate-Q
      'Which person does John like and Mary hate?'

   c. *Dono hito-o*   John-ga   aisitei-te Mary-ga   nikundeiru-no?   (Japanese)
      which person-ACC John-NOM love-and Mary-NOM hate-Q
      'Which person does John love and Mary hate?'

The exact position to which the *wh*-phrases move is not important for our purposes; Cho and
Zhou (1999) analyze this movement as scrambling, and Wu (1999) as *wh*-topicalization. It differs
from canonical *wh*-movement in that it does not induce a weak crossover effect and does not
reconstruct (Wu 1999).

(32) a. Shei$_i$ *ta$_i$* de muqing hen xihuan?   (Chinese)
        who he DE mother very like
        'Who does his mother like?'
     b. Shenme meigeren dou mai-le?   ($\exists$x $\forall$y)
        what   everyone all   buy
        'What did everyone buy?'

   ATB quantifier raising also does not seem to exist, as shown by the lack of a wide scope
reading for the existential quantifier in (33) (see Bošković and Franks 2000 for similar examples).

(33) Every philosopher read *some paper* and every linguist reviewed *some paper*.

   The lack of a wide scope reading for the existentially quantified noun phrase in (33) cannot
be attributed to a more general ban against covert movement out of coordinate structures. Such
movement is generally possible, and it is subject to the Coordinate Structure Constraint (Ruys
1992, Fox 2000, Lin 2002). This is shown by the ungrammaticality of (34a), in which the *wh*-
pronoun moves out of only one conjunct, and by the lack of a wide scope reading for the universal
quantifier in (34b).

(34) a. *I wonder who took *what* from Mary and gave a book to Jeremy.
     b. A student likes every professor and hates the dean.   (*$\forall > \exists$, $\exists > \forall$)

However, as first noted by Ruys (1992), both quantifier raising and covert *wh*-movement may
extract an element out of one conjunct as long as the extracted element establishes a binding
relationship inside both conjuncts.

(35) a. I wonder who took *what* from Mary and gave *it* to Jeremy.
     b. Jeremy will publish *every book* and write *its author* a check.

The contrast in grammaticality between (34a–b) and (35a–b) suggests that the Coordinate Struc-
ture Constraint is a constraint on representations, rather than derivations. Fox (2000), for example,
derives it from a parallelism constraint that requires grammatical constraints to be checked in-
dependently in each conjunct.

If the analysis of overt ATB movement presented in the previous section is on the right track, the lack of covert ATB movement is to be expected. Since movement in ATB structures is crucial for linearization, it has to take place before Spell-Out. Covert movement out of the Parallel Merge structure happens too late in the derivation to be relevant for PF considerations.

4.3 Across-the-Board Wh-*Movement in Languages with Overt Multiple* Wh-*Fronting*

Another interesting consequence of the Parallel Merge account of ATB movement comes from the fact that it clearly shows that ATB *wh*-movement cannot be reduced to a parallelism constraint requiring each conjunct in a coordinate structure to contain a gap. If this were the case, we would expect the structure in (36), where the two conjuncts contain equal numbers of fronted *wh*-phrases and gaps, to be possible.

(36)
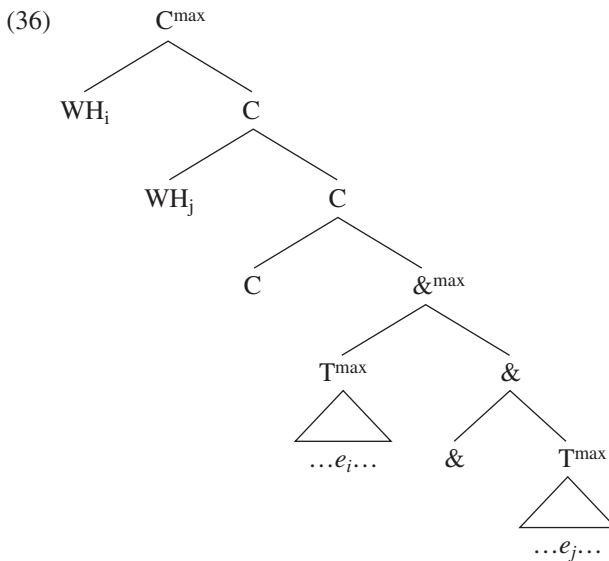
For languages of the English type, the reason why (36) is not available might lie in a ban against multiple specifiers (or whatever is responsible for the lack of overt multiple *wh*-fronting in English). This, however, cannot be the right explanation for languages of the Slavic type, which front all *wh*-phrases overtly in multiple *wh*-questions. However, even in Slavic, ATB *wh*-questions involve a single fronted *wh*-pronoun. This is shown by the ungrammaticality of the Polish example in (37), which contains two fronted *wh*-phrases, one extracted from each conjunct.[17]

---

[17] The ungrammatical status of (37) could perhaps be attributed to a constraint banning sequences of homophonous *wh*-forms. However, exactly the same grammaticality judgments result if one of them is replaced with a nonhomophonous *wh*-form such as the dative *komu*. Furthermore, the ungrammaticality of both (i) and (ii) on page 492 shows that superiority is not a factor, either.

(37) *Kogo$_i$ kogo$_j$ Jan lubi e_i a   Maria kocha e_j?   (Polish)
       whom whom Jan likes   and Maria loves
       'Whom does Jan like and Maria love?'

On the Parallel Merge account, the reason why ATB *wh*-questions contain a single fronted *wh*-pronoun becomes straightforward. There is only a single *wh*-phrase in the premovement structure, which undergoes standard *wh*-movement to a single Spec,C position. In this sense, the Parallel Merge approach takes the magic out of ATB *wh*-movement. There is nothing special about ATB movement per se; what is ''special'' is the structure out of which the *wh*-pronoun moves, namely, the Parallel Merge structure.

4.4 Reconstruction Effects

The Parallel Merge approach to ATB *wh*-questions makes a clear prediction regarding reconstruction effects. Since a *wh*-phrase is shared between two conjuncts in an ATB question, it should reconstruct into both conjuncts simultaneously. The facts here are somewhat inconclusive. With respect to some diagnostics, both conjuncts are affected simultaneously. With respect to others, there is an asymmetry between the first and the second conjunct. This suggests to me that reconstruction diagnostics do not form a unified class, and sometimes linear precedence is also a factor.

Variable binding reconstruction affects both conjuncts. The variable contained in the fronted *wh*-pronoun has to be bound by the quantifiers inside both conjuncts, as shown by the contrast in (38), modeled after Nissenbaum's (2000:44) examples.

(38) a.   Which picture of *his mother* did *every Italian* like and *every Frenchman* dislike?
       b. #?Which picture of *his mother* did *every Italian* like and Mary dislike?
       c. #?Which picture of *his mother* did Mary dislike and *every Italian* like?

Idiom interpretation, strong crossover, and scope reconstruction all point toward the same conclusion. Both conjuncts allow the idiomatic reading of *take pictures*, as shown in (39a–b).

(39) a. Which picture did John take and Bill pose for?
       b. Which picture did John pose for and Bill take?

Strong crossover effects arise within both conjuncts, as shown by the ungrammatical status of both (40a) and (40b).

(40) a. *Whose$_i$ mother did we talk to and he$_i$ never visit?
       b. *Whose$_i$ mother did he$_i$ never visit and we talk to?

(i)   *Kogo$_i$ komu$_j$ Jan lubi e_i a   Maria się   przygląda e_j?   (Polish)
        who.ACC who.DAT Jan likes   and Maria REFL looks.at
        'Who does Jan like and Maria looks at?'

(ii) *Komu$_j$ kogo$_i$ Jan lubi e_i a   Maria się   przygląda e_j?
        who.DAT who.ACC Jan likes   and Maria REFL looks.at
        'Who does Jan like and Maria looks at?'

A further argument in favor of simultaneous reconstruction into both conjuncts comes from the interpretation of *how many* ATB questions including quantifiers. Moltmann (1992:137–138) observes that such questions are multiply ambiguous, as shown in (41).

(41) a. How many books did every student like and every professor dislike?
   b. Seven books.   (*how many* > *&* > *every*)
   c. Student A liked seven books and Prof. B disliked two books; Student C liked nine books and Prof. D disliked four books.   (*&* > *every* > *how many*)
   d. Every student liked seven books and every professor disliked three books.   (*&* > *how many* > *every*)

How many books can have wide scope with respect to both *every* and *and*, in which case the appropriate answer is the one in (41b). It can also have narrow scope with respect to *and* and *every*, in which case a double family-of-questions reading results, given in (41c). Furthermore, it can have narrow scope with respect to *and*, but wide scope with respect to *every*, as shown in (41d). The reading that is crucial for our purposes would be the one in which *how many books* has wide scope with respect to *every* inside the first conjunct but narrow scope with respect to *every* inside the second conjunct (or vice versa), irrespective of its scope with respect to the conjunction. A nice way to bring out the relevant reading (or the lack thereof), suggested to me by an anonymous reviewer, is to replace *every professor* in the second conjunct with *almost every professor*, which obligatorily takes narrow scope.

(42) How many books did every student like and almost every professor dislike?

The prediction that scope affects both conjuncts simultaneously is confirmed. The fact that (43) is not a felicitous answer to (42) suggests that *how many books* cannot have narrow scope with respect to *every student* and wide scope with respect to *almost every professor*.

(43) #Student A liked seven books, and Student B liked three books, and Student C liked nine books, and almost every professor disliked six books.

While variable binding, scope, strong crossover effects, and idiom interpretation show that reconstruction affects both conjuncts simultaneously, anaphor binding shows an asymmetry between the two conjuncts.

(44) a. *Which pictures of himself$_i$ did Mary sell and John$_i$ buy?
   b.   Which picture of himself$_i$ did John$_i$ sell and Mary buy?

Similar asymmetries between the first and the second conjunct arise with respect to Principle C effects and weak crossover effects.[18] Thus, Principle C and weak crossover effects arise only inside the first conjunct, as shown in (45) and (46), respectively.

---

[18] Alan Munn (pers. comm.) points out that the use of resumptive pronouns in Hebrew shows a similar first versus second conjunct asymmetry.

(45) a. *Which picture of John$_i$ did he$_i$ like and Mary dislike?
     b. Which picture of John$_i$ did Mary like and he$_i$ dislike?

(46) a. *Who$_i$ did his$_i$ boss fire and John hire?
     b. Who$_i$ did John hire and his$_i$ boss fire?

At present, I do not have a clear idea of why different reconstruction diagnostics yield different results in that with respect to some diagnostics both conjuncts are equally affected, and with respect to others precedence also seems to be a factor. The important point that I want to emphasize here is that reconstruction *may* affect both conjuncts simultaneously.

5 Conclusion

To conclude briefly, I have argued in this article for a new type of Merge, Parallel Merge, which combines the properties of Internal and External Merge. Parallel Merge creates symmetric structures that become antisymmetric in the course of the derivation. This avoids the linearization issue faced by some of the other multidominance proposals. I have also argued that Parallel Merge is involved in the derivation of ATB *wh*-questions, which accounts for the presence of the so-called matching effects in ATB questions, the lack of covert ATB movement, and the lack of ATB *wh*-questions with multiple fronted *wh*-forms.

References

Borsley, Robert. 1983. A note on the Generalized Left Branch Condition. *Linguistic Inquiry* 14:169–174.
Bošković, Željko, and Steven Franks. 2000. Across-the-board movement and LF. *Lingua* 3:107–129.
Cho, Sungeun, and Xuan Zhou. 1999. The interpretation of *wh*-elements in conjoined *wh*-questions. Ms., State University of New York, Stony Brook.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2001. Beyond explanatory adequacy. MIT Occasional Papers in Linguistics 20. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL.
Chvany, Catherine. 1986. Jakobson's fourth and fifth dimensions: On reconciling the cube model of case meanings with the two-dimensional matrices for case forms. In *Case in Slavic*, ed. by Richard D. Brecht and James S. Levine, 107–129. Columbus, Ohio: Slavica.
Citko, Barbara, and Kleanthes Grohmann. 2000. A new argument in favor of a syntactic focus projection. Paper presented at the GLOW Colloquium, Bilbao.
Collins, Chris. 1988. Conjunction adverbs. Ms., MIT, Cambridge, Mass.
Dyła, Stefan. 1984. Across-the-board dependencies and case in Polish. *Linguistic Inquiry* 15:701–705.
Epstein, Samuel, Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations*. New York: Oxford University Press.
Fox, Danny. 1999. *Economy and semantic interpretation*. Cambridge, Mass.: MIT Press.
Franks, Steven. 1993. On parallelism in across-the-board dependencies. *Linguistic Inquiry* 24:509–529.
Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. New York: Oxford University Press.
Gärtner, Hans-Martin. 1999. Phrase linking meets minimalist syntax. In *Proceedings of the Eighteenth West*

Coast Conference on Formal Linguistics, ed. by Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 159–169. Somerville, Mass.: Cascadilla Press.

Goodall, Grant. 1983. A three-dimensional analysis of coordination. In *CLS 20*, ed. by Amy Chukerman, Mitchell Marks, and John F. Richardson, 146–154. Chicago: University of Chicago, Chicago Linguistic Society.

Goodall, Grant. 1987. *Parallel structures in syntax*. Cambridge: Cambridge University Press.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, Mass.: MIT Press.

Hiraiwa, Ken. 2001. Multiple Agree and the Defective Intervention Constraint in Japanese. In *Proceedings of the 1st HUMIT Conference in Language Research (HUMIT 2000)*, ed. by Ora Matushansky et al., 67–80. MIT Working Papers in Linguistics 40. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL.

Jakobson, Roman. 1958. Morfologičeskie nabljudenija nad slavjanskim skloneniem. In *Selected writings*, vol. 2, 154–183. The Hague: Mouton.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Lin, Vivian. 2002. Coordination and sharing at the interfaces. Doctoral dissertation, MIT, Cambridge, Mass.

Marantz, Alec. 1995. A late note on late insertion. In *Explorations in generative grammar*, ed. by Young-Suk Kim, Byung-Choon Lee, Kyoung-Choon Lee, Hyun-Kwon Yang, and Jong-Yurl Yoon, 396–413. Seoul: Hankuk.

Moltmann, Friederike. 1992. Coordination and comparatives. Doctoral dissertation, MIT, Cambridge, Mass.

Moro, Andrea. 2000. *Dynamic Antisymmetry*. Cambridge, Mass.: MIT Press.

Muadz, Husni. 1991. Coordinate structures: A planar representation. Doctoral dissertation, University of Arizona, Tucson.

Munn, Alan. 1993. Topics in the syntax and semantics of coordinate structures. Doctoral dissertation, University of Maryland, College Park.

Munn, Alan. 1999. On the identity requirement of ATB movement. *Natural Language Semantics* 7:421–425.

Neidle, Carol. 1988. The role of case in Russian syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Nissenbaum, Jonathan. 2000. Investigations of covert phrase movement. Doctoral dissertation, MIT, Cambridge, Mass.

Nunes, Jairo. 1995. The copy theory of movement and linearization of chains in the Minimalist Program. Doctoral dissertation, University of Maryland, College Park.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 355–427. Cambridge, Mass.: MIT Press.

Progovac, Ljiljana. 1998. ''Avoid Conjunction,'' adjunction, and the ''Coordination of Likes'' Constraint. In *Formal Approaches to Slavic Linguistics: The Connecticut meeting 1997*, ed. by Željko Bošković, Steven Franks, and William Snyder, 252–266. Ann Arbor: Michigan Slavic Publications.

Ruys, Eddy. 1992. *The scope of indefinites*. Doctoral dissertation, Utrecht University.

Starke, Michal. 2001. Move dissolves into Merge: A theory of locality. Doctoral dissertation, University of Geneva.

Wilder, Chris. 1999. Right node raising and the LCA. In *Proceedings of the Eighteenth West Coast Conference on Formal Linguistics*, ed. by Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 586–598. Somerville, Mass.: Cascadilla Press.

Williams, Edwin. 1978. Across-the-board rule application. *Linguistic Inquiry* 9:31–43.

Wu, Jianxin. 1999. A minimalist account of quantification in Chinese. Doctoral dissertation, University of Maryland, College Park.

Zhang, Niina. 2002. Move is Remerge. Paper presented at GLOW in Asia 3.

Zoerner, Edward Cyril. 1995. Coordination: The syntax of &P. Doctoral dissertation, University of California, Irvine.

Department of Linguistics
University of Washington
Box 354340
Seattle, Washington 98195-4340

bcitko@u.washington.edu